DOCUMENT RESUME

ED 077 115 EA 005 041

AUTHOR Jacobs, James W.
TITLE Releasing Test Scores: Urgent or Unthinkable?
PUB DATE 27 Feb 73
NOTE 11p.; Paper presented at American Association of School Administrators Annual Convention (105th, Atlantic City, New Jersey, February 24-28, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Confidentiality; *Intelligence Tests; Parent Conferences; Report Cards; *Reports; Speeches; *Standardized Tests; Test Interpretation; *Test Results

ABSTRACT

In this paper, the author reviews literature on the subject of releasing test scores and discusses constraints (in terms of test scores released and target groups receiving test scores) which he believes should govern such releases. The author suggests two phases for implementing release of test scores. One phase would involve parent conferences, districtwide appraisal programs of intelligence, districtwide appraisal in selected subject fields, and the appraisal data reported to parents by report cards. The other phase would be directed toward releasing test scores in relationship to salient characterists of the school system (teacher-pupil ratics, staff turnover statistics,etc.) that affect the learning of its clientele. The author expresses his preference for a community-school profile over the release of scores on a systemwide or a school-by-school basis. (JF)

ED 077115

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

AMERICAN ASSOCIATION OF SCHOOL ADMINISTRATORS
105th Annual Convention
Atlantic City, N. J. February 24-28, 1973

SPEAKER: James W. Jacobs, Associate Superintendent, Office of Planning, Management, and Computer Services, Montgomery Cty Sch. Rockville, Md.

TOPIC: RELEASING TEST SCORES: URGENT OR UNTHINKABLE?

PLACE: Coral Reef Room, Shelburne Hotel

TIME: Tuesday, February 27, 2:30 P.M.

PROGRAM: Page 79

FOR RELEASE UPON DELIVERY

If the purpose of your visit to this meeting is to hear the speaker justify your position that releasing test scores to the public for the schools in your school district is urgent, essential and we ought to get on with it - then you are going to be disappointed. If, on the other hand, the purpose of your visit is to hear the speaker defend your position that releasing test scores to the public for schools in your school district is unthinkable, inane and your board of education shouldn't even discuss it - then you are going to be disappointed too! To put it bluntly, it may well be that we all go away disappointed and keep in mind, except for the panel - I'm the only one in the room that has read the speech! The reason I am stressing the idea that this meeting may be a disappointment is due to the nature of the topic which is to hold our attention for the next hour or so! The statement of the problem - Releasing Test Scores: Urgent or Unthinkable - is pure unadulterated over-simplification. As any good educator can tell you - it just isn't that simple! Frankly, I have no problem with releasing test scores as long as the test scores released and the target group which receives these scores fit the constraints which I feel should govern the release program. Without beating a dead horse on this - suffice it to say it is not a yes or no question!

When I was asked by the AASA home office to accept this assignment, I had a queezy feeling in my stomach as I signed and returned the acknowledgement form. Before you discover them for yourself, permit me to state the facts: (1) I am not an expert in the field of pupil and program appraisal, and (2) I am not a research or statistical analyst. Consequently, I will not dabble in these areas of specialty today, but rather I will (1) attempt to share with you the results of my brief sojourn into the recent periodical literature pertaining to the general topic and (2) outline some of the conditions and considerations which might be food for thought as you return to your school district to face the issue of releasing test scores this year, next year, some year or perhaps - not while I'm Superintendent!!

-more-

EA 005 041

FILMED FROM BEST AVAILABLE COPY

Even though my original queezy feeling is still with me because so many of you in this room today are experts in the topic area, I must admit that I have enjoyed reviewing the literature and getting a flavor of what is occurring in some school districts relating to how test scores are disseminated external to the district.

Though I had not, many of you have probably read the article in the April, 1972 issue of Nations Schools entitled, "Releasing Test Scores: Urgent or Unthinkable." The author, Mr. Gene R. Hawes is a free-lance writer and former editor for the College Entrance Examination Boards and is the author of Education Testing for the Millions: What Tests Really Mean for your Child. I have not read his book, but I found Mr. Hawes' article quite informative. It offers to the reader not only a capsule summary of the status of releasing test scores to the public, but cites several school districts which have developed exemplary programs. It could even be retitled, "Everything I Always Wanted to Know About Releasing Test Scores but"- and you know the rest! For those who really need to know, Mr. Hawes offers to the reader twelve sound ways to announce test results. (Now that I have read the article I am aware of one other person who must have read it - the AASA staff member who is in charge of developing topics for this conference!)

Testing has been making the scene in education for a good many years. In fact, since World War I schools have, in part, been sorting and training its manpower through the help of standardized tests. Marjorie C. Kirkland in the October 1971 issue of Review of Educational Research, in an article entitled, "The Effects of Tests on Students and Schools" states:

> Although the use of tests has increased rapidly throughout this century, it has been particularly pronounced in the last 15 years in the public schools . . .In view of the facts that there were approximately 45 million elementary and secondary school children during this period, this represents three to five standardized tests per pupil per year.

The Kirkland article goes on to state:

> Testing is indeed big business. To supply the demand for tests, there are 150 organizations whose major business (or at least an important part of it) is the development and publication of tests.

She continues:

> The ways in which these tests are used in schools multiply each year. They are being used to assess more and more aspects of behavior and, to an increasing extent, to make practical decisions about the individual, decisions of vital importance for both the individual and society. Illustrative of test uses are: (a) grouping pupils within a class for instructional purposes, (b) assigning pupils to classrooms, (c) placing new pupils, (d) identifying pupils who need special diagnostic study and remedial instruction (e) helping a pupil to select courses and subjects and to solve personal and social problems, (f) helping a pupil to set educational and vocational goals,

(g) providing information to parents, community, and outside agencies, and (h) evaluating the curricula and the efficiency of the school. The use of standardized tests that is perhaps most visible to the public is for admission to higher institutions of learning.

Of these nine uses of tests, note that seven are allied directly to the operation of the school; one is directed toward helping span the high school/higher education gap and only one use relates to providing information to the community.

In the September issue of Today's Education, Lillian Zach in an article entitled, "The IQ Debate" brings another element of the issue to focus in this statement:

> Herein lies a dilemma: Intelligence was only vaguely defined by the test maker, but the tests were used to define intelligence. This is perhaps the greatest failure of the testing movement in the United States. The pragmatic value of the mental test is undiminished. Test scores are good indicators of functioning abilities as long as their limitations are clearly understood, but these test scores should not be used outside of their immediate significance. The failure lies not in the mental tests themselves, but in the perversion of the test results by investigators and social philosophers who use numbers in support of far-reaching positions. It is unfair both to the person tested and to the test itself to say that the scores of any one individual represent support for broad statements concerning human development.

Professor Zach goes on to say -

> There is nothing inherently wrong with practical definitions as long as they are clearly understood. The tests, after all, were developed to measure those aspects of human behavior which correlate well with scholastic achievement. In order to succeed in school, an individual must demonstrate certain types of abilities. If we develop tests to measure these abilities and if they prove to be valid and reliable instruments, we are measuring some form of intellectual ability. But if we lose sight of what we are measuring and if we claim for the test qualities for which it was never intended, we can be led into invalid implications.

I found the Zach article to be especially interesting in the following excerpts:

> Another assumption is that the mental test is a sampling of behaviors which directly reflect the general capacity for learning. Actually, all available intelligence tests are direct measures only of achievement in learning. We wrongly equate the inferences from scores on IQ tests to some native inherent trait. Many persons think of intelligence as a discrete dimension existing within the individual and believe that different people have different amounts of it.
>
> In a certain sense this is true, but one's intelligence is not a characteristic of a person so much as it is a characteristic of the person's behavior. We can only hope to measure or observe manifestations of it.

-more-

It is also not possible to add up the elements of someone's intelligence in the same way that you can count the number of fingers on his hand. Although two people can have the same IQ score, they may demonstrate quite different abilities by virtue of the fact that they succeeded on different parts of the test. All too often, undue weight is given to an IQ score, although numerical assignment of a child to a man-made concept, untied to real characteristics of the child, tells us very little. Even more unfortunate, parents and some teachers are led to believe the IQ concept has deeper significance than its meaning as a score.

Richard R. DeBlassie, in his article "Test Anxiety: Education's Hang-up" in the May, 1972 issue of the Clearing House explores relevant studies on test anxiety and suggests approaches which might minimize this reaction. The last section of the article speaks to the question,

What does the nature of test anxiety imply for educational goals and practices? It is in the treatment of the scores and the interpretation of their meanings that real problems arise. Effective test interpretation, including both standardized and teacher-made tests, will result in the accumulation of experiences by the various staff members that will be fed back into staff discussions of the school's testing program. There should be a continuing opportunity for all school personnel and parents to report on their experiences with the tests that have been used. This feedback should also place a great deal of emphasis on the students' reactions to their testing experience. It is essential that the reactions of children which give evidence of anxiety, pressure, or fear in relation to tests be carefully noted and considered, especially when test results are interpreted and used for instructional, guidance, and administrative purposes.

The criticisms leveled at external testing programs by the 1962 Joint Committee on Testing (AASA-NEA) warrant consideration, additionally, in terms of the use and interpretation of test results. It is of great importance that finality not be attached to test scores. They should be conceived as of indices of a sample behavior taken at a specific point in time and subject to the moods of the individual taking the test. Test users should realize that the evaluation of the pupil through tests is limited and that tests represent only one of many techniques for pupil appraisal.

Effective use of test results is best seen as an aspect of the schools' regular program of conferences between teachers and parents, children and teachers, counselors and children or parents.

My review of recent periodical literature uncovered some interesting comments concerning what use can be made of pupil records. Thomas W. George, NASSP assistant for Legal and Legislative Services writes in the September 1972 issue of the NASSP Bulletin that:

- The dominant issue to emerge is the legal character of pupil records. Whether student records are private or public has not been clearly decided by the courts.

Mr. George calls attention to the question:

Are Student Test Scores a 'Record?'

He notes that -

> A decision by New Jersey's Commissioner of Education (Citizens for Better Education vs. Board of Education of the City of Camden, December 20, 1971) held student test scores not to be a part of a student's record. The case grew out of a citizen's demand that the board make public the results of standardized achievement tests. The citizen claimed that the test scores were a matter of public interest and concern and that their release would enable the citizens to determine the educational achievement of the students and efficiency of the education program. Although the commissioner agreed that the test scores were part of the student's permanent record, and that the permanent record by state law is open to persons with legitimate interest, he nevertheless ruled that since the testing program was not required by state law, the board retained the authority to release or withhold this information at its discretion.

As a note of caution, Mr. George goes on to say,

> But this is not the kind of decision upon which totally to base one's school policies.

In this month's issue of Nations Schools, you may recall in the "School Law" section by Lawrence W. Knowles this statement:

> Releasing standardized test scores violates the first and fourteenth Amendments and denies children 'their right not to be stigmatized by degrading titles.' Or so says the New Jersey Education Association and a group of parents who filed suit to prevent publishing results of tests given to fourth and twelfth graders recently. Named as defendant is State Education Commissioner Carl L. Marburger.

It appears to me that, as usual, the Commissioner is caught in the middle!

One other item from the literature which seems pertinent is an item included in the December 11, 1972 issue of Education, USA. The item, "National Assessment Moves in New Directions." indicates that:

> Twenty-five states have adopted, or are considering adopting, a a localized version of 'national assessment' to evaluate student achievement. The National Assessment of Educational Progress (NAEP), a federally funded program of the Education Commission of the States, originally did not intend its tests to be adopted by states. But educators, who at first opposed the idea of state/national comparisons, are now calling for 'more information to help us assess our schools,' says staff director J. Stanley Ahmann. Sixteen states are actively working to adopt the testing questions and techniques, and nine others are considering it. Maine has just

released the results of its testing survey, a project called by Anmann 'the most complete to date.' Maine State Senator Bennett D. Katz says his state's legislature has found the tests to be 'the first effective tool we have had to answer the question, 'What are we buying for the ever-increasing dollar demand?'

Iowa, Massachusetts and Connecticut have also begun actual testing.

NAEP is also moving to expand its project in two other ways: it plans to begin interpreting its testing data and collecting more background information on students to determine what factors influence pupil achievement. The first in-depth interpretation of the NAEP data is being done by the National Science Teachers Association, which will analyze NAEP science findings and consider their significance for education and teaching. NAEP has also approved proposals by the Social Science Education Consortium and the National Council for the Social Studies to interpret social studies and citizenship results. Interest in working with NAEP has also been expressed by the Association for Supervision and Curriculum Development, the International Reading Association and the National Council of Teachers of Mathematics. The next step, says Ahmann, will be to involve educators in setting up desired levels of performance in subject areas. The NAEP steering committee has also accepted a proposal to expand the project's research into factors which may affect a child's learning. NAEP now collects data on sex, race, age, level of parents' education and community type and size, but Ahmann hopes NAEP eventually will collect information on '20 factors or more,' including the child's intelligence, his need for achievement, the economic level of his parents and the ethnic mix of his community. This kind of analysis, says Ahmann, 'could lead to a major revision of NAEP.'

-more-

As noted earlier in my remarks, by far the most comprehensive coverage of the topic in recent periodical literature is the article by Gene R. Hawes. Mr. Hawes draws attention to the fact that in growing numbers, school officials are releasing test scores for their school districts and some even release test scores by individual schools. He goes on to state that the release of test scores by individual schools was once considered very inflammatory, and it is still considered so by some officials. The reason, Mr. Hawes cites, for releasing test scores to the public is that the schools must satisfy the increasing needs and demands of the public for accountability and for hard data on what schools are accomplishing with public monies.

As part of his doctoral study, Lyle O. Young, in May 1972, surveyed the eighty largest public school systems in the nation to determine the current practices of large public school systems in the release of achievement test results. This included the systems with pupil memberships of 50,000 or more. Seventy, or 88 percent,of the systems returned completed questionnaires.

Thirty-two (46%) of the seventy school systems did not release any 1970-71 achievement test results. Eighteen (26%) of the systems released only system-wide averages, while twenty (28%) released school-by-school results in addition to the system-wide results.

Young states in the summary of findings:

> The dilemma school boards and superintendents face is further illustrated by the responses to an item concerning the policy to be adopted with the 1971-72 test results. Six systems do not plan to follow the same policy with their 1971-72 results. ...(six) plan to release both the system and school-by-school averages for 1971-72. ...(one) had released 1968-69 and 1969-70 school-by-school results but did not do so during the year of this survey because of a financial bind. ...(one) released school-by-school results in 1969-70, but plans to follow an alternate year release policy due to the limited variation from year to year. Also, (one) plans to release school-by-school results of the 1972-73 results.

In addition, six more school systems indicated that they were uncertain as to the 1971-72 policy.

Young goes on to state in the summary:

> The techniques utilized in reporting test results on a school-by-school basis vary considerably. The reports range from a simple list of schools with the results listed for one grade only, to a comprehensive one-thousand page report listing the achievement test results as well as related data. Most systems release accompanying data about the community served by the school, and related school factors such as class size and staff qualifications or experience.

This survey included two items which sought information about the reasons and disadvantages of the release policies of the schools. The reason usually stated for withholding school-by-school results was to avoid comparisons among the schools and to avoid confusion and misinterpretations. The major disadvantage in not releasing results was that the public has lost confidence in the schools because of the policy of confidentiality.

The major reason cited for releasing test results wasthat "the public has the right to know." Several respondents, Young states, indicated that they released results because of extreme pressure their boards had received. Others felt that "information is better than speculation." The respondents felt that it was good public relations; it increased community involvement; and it improved community understanding. The disadvantages were that test results were often misinterpreted by the public and misinterpreted by the press. Militant groups sometime use the results to "show" that the quality of education in minority areas is inferior. Staff morale was sometimes bad in low-achieving schools because of the comparisons made among schools.

Young concludes the summary with this comment:

> An item concerning the future of testing and the disclosure of results indicated that most respondents felt that the release to the public of school-by-school results would increase. The accountability movement, which demands 'proof' that children are learning has provided this impetus.

(Parenthetically, I might comment to Mr. Young that the reason our school system did not respond is that we were too busy compiling our annual test report.)

The accountability movement seems to be sweeping like wildfire throughout education. It is almost as if in the past we have been unaccountable. My concept of accountability may be a little different than yours. First, accountability tends to convey an after-the-fact action; somewhat like accounting for dollars already expended and then wondering why you overspent your budget. Or, somewhat like locking the barn door after the horse has run away. Certainly any good financial system has an encumberance phase which relates dollars available for expenditure to the anticipated expenditure. Building in an encumberance capability helps discipline the program manager to stay within appropriated funds. So it is with my accountability model which I prefer to call a quality assurance model. If I were to conceptualize my quality assurance model, I would include the entire educative process within the model and establish a monitoring or continuous progress system which would alert persons responsible for specific program areas to where quality control or assurance is lacking, has diminished, or is not under control. Although it may sound inhumane the model which I envision would have its primary focus on the raw material - the learner. The model would be activated at the time of initial learner contact with the school system. A tracking system would be required and each learner would be monitored from entry to withdrawal. The constraint of time this afternoon prohibits a detailed analysis of my model. For that I am somewhat thankful since I have not fully

conceptualized, and documented, much less implemented a working model. However, one essential and vital component of the model is a comprehensive school information system and I am pleased to report that in our school district, we have conceptualized and initiated the development and implementation of this type of data base. But that is another story!

In his article Hawes calls attention to several school districts which have released test scores to the public. Since there is a big difference, it must be noted that some of the systems release only district-wide test scores while others have chosen to give them school by school. In this later category are Tulsa, Oklahoma; Columbus, Ohio; and New Rochelle, New York. California districts have a mandate from their state to release district-wide test scores with some districts opting to release on a school-by-school basis.

At this point it might be interesting to look at both sides of the school-by-school test score release program when related to various levels of accountability.

Since I am quite sure the release of test scores by school issue will not be resolved quickly, I urge you to consider both sides of various considerations and develop a conclusion for each consideration which is in the best interest of your school district and not developed on the basis of what some other school system is doing.

It is not uncommon for a school system to present a system-wide test report to its Board of Education in open session. In my opinion this is the minimum effort which every school district should implement this year. I spoke a moment ago about an accountability model or, I should say, quality assurance model. Another element of the model would be the "reporting of pupil progress module". I would construct this module in two phases. Phase I would incorporate elements which can be implemented earlier and easier than others or are routines that already may be practices in many districts:

1. Parent conferences on a regular cycle for all children K-12.

2. A pupil report card system which includes several alternatives to a so-called standard marking system, i.e. A-B-C.

3. A district-wide appraisal program of intelligence, basic skills and academic progress tests for grades 3-5-7-9-11.

4. A district-wide appraisal program in selected subject fields for grades 4-6-7-8-9 and 11.

5. The reporting of appraisal data via:

   A. A teacher developed report card for parents at least four times each year, preferably six.

   B. A test report to parents following the administration of all standardized tests.

C. An individual pupil test card for each school with a copy for the teacher of the student when the test was administered and a copy for the teacher of the student the following year.

D. Sticky labels for permanent records cards in each school.

E. A district-wide annual Test Report to the Board of Education presented and discussed in public session with a news release to the local media.

6. An annual program report by each school to its community. This report would not only inform parents of test results but would, in addition, analyze each program in the school.

Phase II of my model would be directed toward releasing test scores in relationship to salient characteristics of the school system that affect the learning of its clientele. Rather than releasing test scores on a system-wide basis or school-by-school in isolation, I believe it is important to provide a community-school profile in both situations. Fully do I recognize that establishing and maintaining an adequate and proper data base for profile development is a long-range goal for most, if not all, districts, especially the large districts which require a computer based information system. Assuming the data base, what are its major components? Certainly it would include the minimal capability to retrieve and relate:

1. teacher/pupil ratios and class size statistics
2. adult professional and non-professional/pupil ratios
3. staff turnover statistics
4. pupil mobility data
5. minority enrollment figures or ethnic mix
6. attendance characteristics
7. cost by program and by school
8. personnel data (training, certification, experience, etc.)
9. demographic information for the community
10. economic data for the community
11. pupil personnel data

Even though the school district releases test scores as part of a composite profile for each school and/or the school district there is no assurance that the results of the testing program will have greater meaning. Hawes states it this way:

> Because no satisfactory systematic method of connecting test results and input factors has been developed, however, avoid interpretation. Leave it to the reader to connect specific factors with specific results intuitively.

Although this statement must be taken seriously, it does not mean that the school district hides its head in the sand. Rather, the school district should provide a comprehensive interpretation of the broad significance of the test results as well as an interpretation of each area potentially fraught with controversy.

Because we have accidents in our transportation program there are no plans to my knowledge, to give up travel. Because we have problems in testing programs as well as reporting their results, I see no reason for throwing the baby out with the wash. It seems to me that there is a joint responsibility here. The test industry and education and education and the public must work together more closely, agree on objectives and resolve on a mutual, cooperative basis many of the horror stories with which some districts deal today. Certainly, our mission in education is needed, our results must be known! Let's face up to the challenge and work toward assuring a quality program for each student.

May I leave you with two resolves:

1. to establish a program for reporting of appraisal data to your community; and

2. to face this issue of releasing appraisal data logically rather than emotionally.

---ooo---